

Comparando dois coeficientes angulares em amostras independentes

Para verificar se dois coeficientes angulares de duas equações de regressão (provenientes de amostras diferentes) são iguais entre si, adotaremos as seguintes hipóteses:

H0: coeficiente angular₁ = coeficiente angular₂

Ha: coeficiente angular₁ ≠ coeficiente angular₂

A estatística do teste será dada por:

$$t = \frac{a_1 - a_2}{\sqrt{S_{a_1}^2 + S_{a_2}^2}} \sim T(n_1 + n_2 - 4)$$

Tal como o teste t para duas amostras independentes, podemos assumir dois cenários ao se calcular a estatística do teste: podemos assumir que a variância das duas amostras **é igual** ou **não**.

Regra do polegar: se a amostra maior apresentar desvio padrão mais de duas vezes maior que o desvio padrão da menor amostra, assuma que as variâncias não são iguais.

Consideremos um desdobramento do exemplo dado em aula, da quantidade de cigarros consumidos por dia como preditor da expectativa de vida. Entretanto, desta vez, teremos duas amostras: indivíduos do sexo masculino e do sexo feminino (Acompanhe no arquivo **Comparação_Coeficientes_Alunos.xls**)

Homens	
Cigarros (x)	Exp vida (y)
5	80
23	78
25	60
48	53
17	85
8	84
4	73
26	79
11	81
19	75
14	68
35	72
29	58
4	92
23	65

Mulheres	
Cigarros (x)	Exp vida (y)
22	88
7	95
20	86
23	60
15	82
34	75
4	80
40	68
8	93
16	77
11	72
52	67
3	90
31	66
18	72
8	78

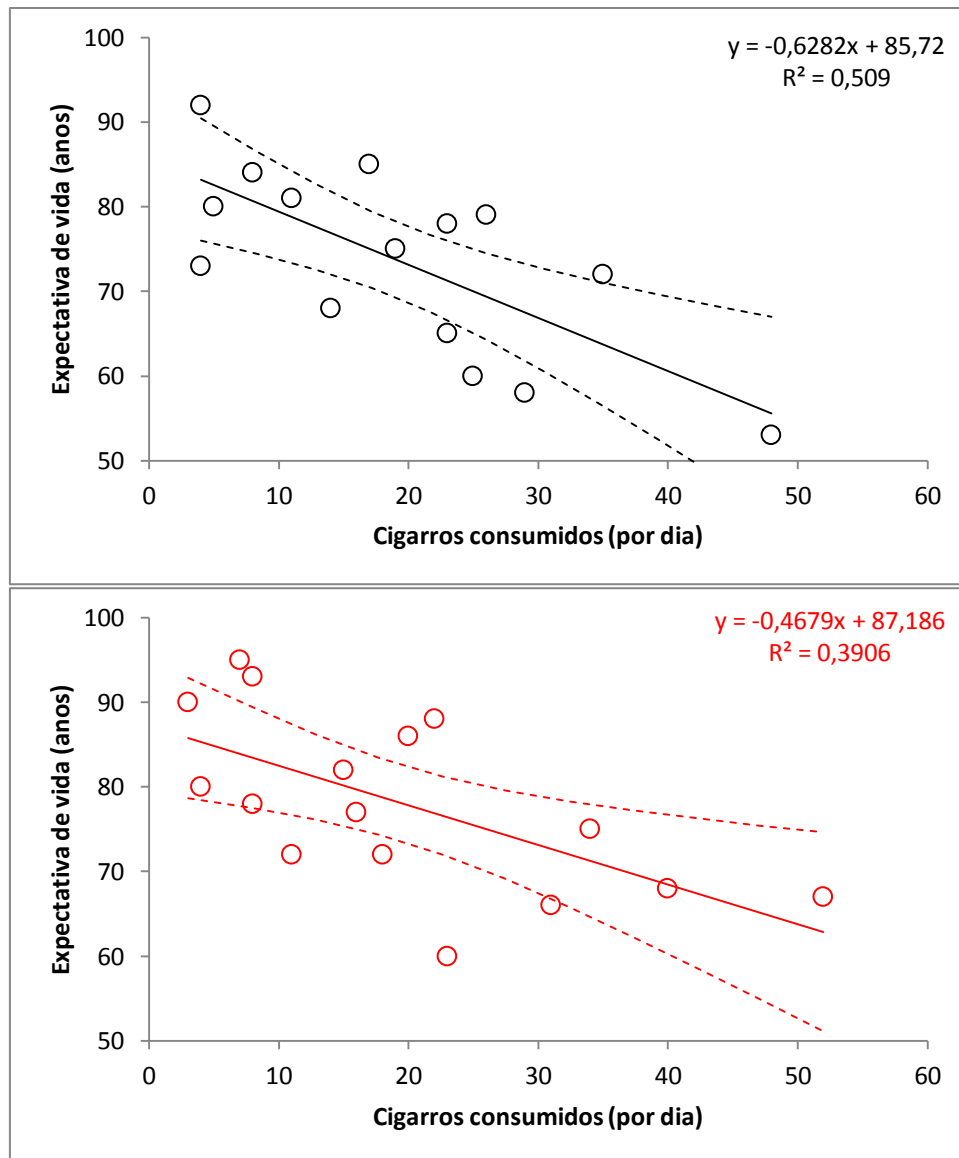


Figura 1. Gráficos de dispersão dos exemplos. Avaliação do gráfico de dispersão com linhas de tendência para amostra de homens (acima) e mulheres (abaixo).

Conforme evidenciamos em aula, podemos avaliar cada conjunto de dados individualmente (homens e mulheres, separadamente) e verificar se o modelo de regressão é ou não um modelo preditivo melhor que a média, se os coeficientes angulares de cada equação são ou não diferentes de zero, avaliando os pressupostos e índices de qualidade do modelo.

A pergunta que resta é: será que existem evidências suficientes para afirmar que os coeficientes angulares das duas equações de regressão são diferentes entre si?

	Homens	Mulheres				
N	15	16				
a	-0,628	-0,468				
EP _{xy}	7,975	8,379				
DP _x	12,454	13,852				
EP _a	0,171	0,156				
Unpooled - unequal variance case			Pooled - equal variance assumed			
EP _{b1-b2}	0,232		Var	67,026		
t	-0,692		EP _{b1-b2}	0,233		
gl	27		t	-0,689		
t-crit	2,052		gl	27		
alfa	0,05		t-crit	2,052		
p	0,495		alfa	0,05		
Sig	NÃO		p	0,497		
			Sig	NÃO		

Homens	gl	SQE	MQ	F	Sig
<i>Regressão</i>	1	856,99	856,99	13,476	0,0028
<i>Resíduo</i>	13	826,74	63,60		
<i>Total</i>	14	1683,73			

Mulheres	gl	SQE	MQ	F	Sig
<i>Regressão</i>	1	629,97	629,97	8,972	0,0096
<i>Resíduo</i>	14	982,96	70,21		
<i>Total</i>	15	1612,94			

Conforme podemos observar a partir dos cálculos acima, assumindo-se ou não as variâncias iguais, a hipótese nula não pode ser rejeitada neste exemplo, e, portanto, conclui-se que não há diferença significativa entre homens e mulheres na expectativa de vida em função da quantidade de cigarros consumidos ($t_{(27)} = -0,689$, $p = 0,497$). Para ambos os gêneros, entretanto, o modelo foi significativo ($F_{(1,13)} = 13,476$, $p = 0,003$ e $F_{(1,14)} = 8,972$, $p = 0,010$, para homens e mulheres, respectivamente), mas as inclinações das retas não são suficientemente distintas entre si, conforme podemos evidenciar pela sobreposição dos dados em mesmo gráfico:

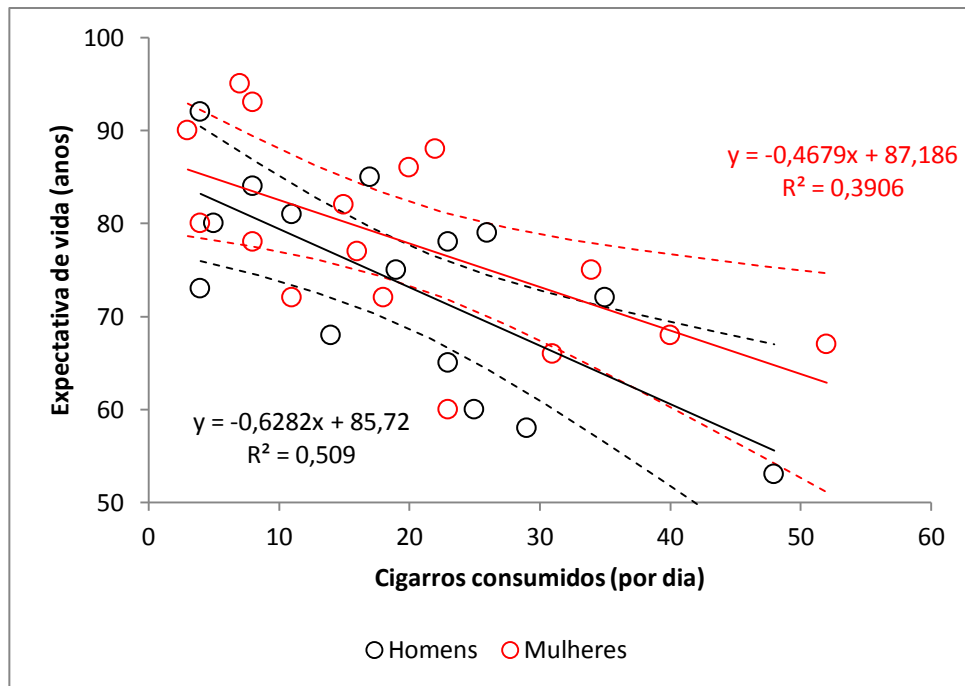


Figura 2. Sobreposição dos dados em gráfico de dispersão. Conforme evidenciamos, não existem diferenças significativas entre os coeficientes angulares dos dois modelos ($t_{(27)} = -0,689$, $p = 0,497$).